Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects

Bin Shen[1,4], Wensheng Zhang[2,4], Jun Zhang[1,4], Jiankui Zhou[1], Jianying Wang[1], Li Chen[1], Lu Wang[3], Alex Hodgkins[2], Vivek Iyer[2], Xingxu Huang[1] & William C Skarnes[2]

Bacterial RNA–directed Cas9 endonuclease is a versatile tool for site-specific genome modification in eukaryotes. Co-microinjection of mouse embryos with *Cas9* mRNA and single guide RNAs induces on-target and off-target mutations that are transmissible to offspring. However, Cas9 nickase can be used to efficiently mutate genes without detectable damage at known off-target sites. This method is applicable for genome editing of any model organism and minimizes confounding problems of off-target mutations.

Zinc-finger nucleases and transcription activator–like effector nucleases have proved effective for genome modification in many model systems, but their use is limited by the need to engineer specific proteins for each target site[1,2]. Use of the clustered, regularly interspaced, short palindromic repeats (CRISPRs) and Cas9 endonuclease, first described in bacterial and archaeal immunity[3], is a simpler and more versatile approach to engineer the eukaryote genome. The DNA-cleaving activity of CRISPR-Cas9 can be programmed with single guide RNAs (sgRNAs) that recognize specific DNA sequences[4–14]. Introduction of multiple sgRNAs provides a way to modify multiple genes at once[8,12–14]. The initial enthusiasm for CRISPR-Cas9 technology has been tempered by recent studies in cultured cells that show substantial off-target damage at related sites[15–18]. Up to five mismatches are tolerated in the sgRNA sequence, which translates into many hundreds of potential off-target sites in the genome. First reported for zinc finger nucleases[19,20], nickase versions of the Cas9 protein directed to paired sites may offer improved specificity[17,21]. In this study, we document co-inheritance of on-target and off-target mutations after injections of one-cell mouse embryos with *Cas9* mRNA and sgRNAs and describe a method for using Cas9 nickases for *in vivo* genome editing that minimizes off-target damage to the genome.

Previously, we had demonstrated efficient *in vivo* modification of a *GFP* transgene from co-injections of one-cell embryos with *Cas9* mRNA and a single sgRNA[11]. Encouraged by our success, we mutated the X-linked androgen receptor (*Ar*) gene *in vivo* and investigated whether *Ar* mutations and off-target damage co-occur in founder animals and their offspring. We constructed two sgRNAs that bind adjacent sites on opposite strands in exon 1 of the *Ar* gene. After we co-injected 79 embryos with *Cas9* mRNA and the two sgRNAs, we recovered 20 pups from two litters. We extracted genomic DNA from the tails of the founder mice and subjected products from PCR amplification of the target region to the T7 endonuclease 1 (T7EN) cleavage assay[11] (**Supplementary Fig. 1a**). We detected cleavage products in five founder mice (founder 5 (F5), F10, F11, F18 and F20), and we confirmed the mutant *Ar* alleles present in the tail samples by sequencing (**Supplementary Fig. 1b** and **Supplementary Table 1**). Similar to reports in mouse[11,13,14] and in zebrafish[7], introduction of *Cas9* and sgRNA induced a range of deletion sizes (5–79 base pairs (bp)), and two of four female founders (F10 and F18) carried no wild-type alleles, which suggested biallelic targeting. The single-mutant male founder mouse (F20) carried four different mutant *Ar* alleles and displayed a feminized phenotype typical of androgen receptor mutants in mice and in humans[22]. These data confirm efficient *in vivo* genome modification of the somatic compartment by injection of one-cell embryos with *Cas9* mRNA and sgRNAs and provide additional evidence for continued nuclease activity in early cleavage-stage embryos[23].

Next, we looked for off-target damage in *Ar* mutant founder mice. We PCR-amplified eight related genomic sites for each sgRNA (AR-A and AR-B) from three mutant founders (F10, F18 and F20) and subjected the amplicons to the T7EN cleavage assay (**Supplementary Fig. 1c**,**d**). We detected cleavage bands for three off-target sites (OTS1, OTS3 and OTS9) and validated off-target damage by sequencing (**Supplementary Table 2**). To test whether both *Ar* and off-target mutations are transmitted through the germ line, we bred two mutant female founders (F10 and F18) and genotyped F_1 offspring. We observed transmission of mutant *Ar* alleles to offspring at the expected Mendelian frequency (**Supplementary Fig. 2a**,**b**). Female F10 transmitted one of two *Ar* deletion alleles to a single male offspring. This male did not show a feminized phenotype, but the transmitted allele was an in-frame deletion of 39 bp. Mouse F18 transmitted two mutant alleles to male offspring: one in-frame deletion of 78 bp and one out-of-frame deletion of 58 bp. Only the male offspring (designated F18-1) with the out-of-frame mutation displayed a feminized phenotype (**Supplementary Fig. 2c**), whereas males with the 78-bp in-frame deletion appeared to be normal.

[1]Ministry of Education Key Laboratory of Model Animal for Disease Study, Model Animal Research Center of Nanjing University, Nanjing, China. [2]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. [3]Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China. [4]These authors contributed equally to this work. Correspondence should be addressed to X.H. (xingxuhuang@mail.nju.edu.cn) or W.C.S. (skarnes@sanger.ac.uk).

RECEIVED 26 AUGUST 2013; ACCEPTED 24 JANUARY 2014; PUBLISHED ONLINE 2 MARCH 2014; DOI:10.1038/NMETH.2857

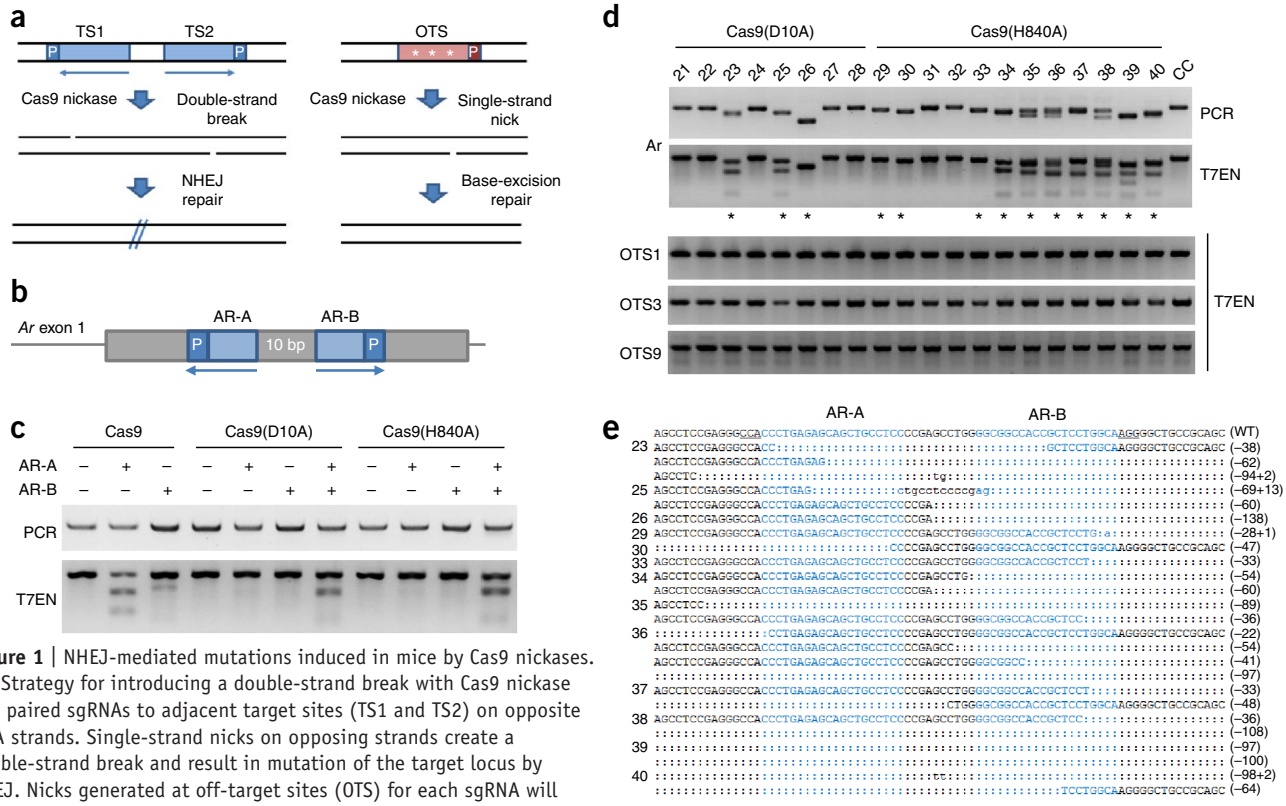

Figure 1 | NHEJ-mediated mutations induced in mice by Cas9 nickases. (**a**) Strategy for introducing a double-strand break with Cas9 nickase and paired sgRNAs to adjacent target sites (TS1 and TS2) on opposite DNA strands. Single-strand nicks on opposing strands create a double-strand break and result in mutation of the target locus by NHEJ. Nicks generated at off-target sites (OTS) for each sgRNA will be corrected by the base-excision repair pathway. P, PAM site. (**b**) Arrangement of Cas9 sites in exon 1 of the mouse *Ar* gene. (**c**) Indels in *Ar* of mouse NIH3T3 fibroblasts with the indicated combinations of Cas9 nuclease or nickases (D10A or H840A variants) and sgRNAs (AR-A, AR-B), detected by the T7EN cleavage assay of PCR products amplified from cells. (**d**) *In vivo* mutation of *Ar* by Cas9 nickases detected in 20 founder mice (F21–F40). PCR products were subjected to the T7EN cleavage assay. CC, F_1(CBA × C57Bl/6) control sample. Bottom gels show assay for cleavage at indicated OTSs for the AR-A sgRNA. (**e**) Sequence of 25 *Ar* mutant alleles present in 13 of the 20 founder animals from co-injections of Cas9 nickases and sgRNAs, AR-A and AR-B target sites in blue. PAM sites are underlined. Inserted sequences are in lower case.

Finally, the single male offspring of the F10 female co-inherited one of two alleles of the off-target site OTS3 (**Supplementary Fig. 2d**,**e**). Our results demonstrate that sgRNA–Cas9 endonuclease can efficiently induce both on-target and off-target damage in founder mice and both on-target and off-target alleles may be co-inherited by their progeny.

We explored the use of CRISPR-Cas9 nickases[4–6] for *in vivo* genome editing with minimal off-target damage (**Fig. 1a**). Because nicked genomic DNA is corrected by the endogenous base-excision repair pathway[24], Cas9 nickases would be expected to induce little or no damage to the genome[5,6]. However, if single-strand nicks are located close together on opposite DNA strands, then the resulting double-strand break may be imprecisely repaired by nonhomologous end joining (NHEJ). This principle has been originally reported for zinc finger nickases[19,20] and more recently for Cas9 nickase[17,21].

We first co-transfected mouse fibroblast cells with plasmids encoding Cas9(D10A) or Cas9(H840A) nickases and *U6*-driven expression plasmids for AR-A and AR-B sgRNA. The AR-A and AR-B target sites, separated by 10 bp, are oriented in a tail-to-tail configuration in exon 1 of *Ar* (**Fig. 1b**). Cleavage 3 bp upstream of the protospacer adjacent motif (PAM) site by Cas9(D10A) nickase is expected to create a double-strand break with 5′ overhangs, whereas Cas9(H840A) should generate 3′ overhangs. We detected modification of the *Ar* gene by the T7EN cleavage assay in cells transfected with Cas9(D10A) or Cas9(H840A) vectors, but only in the presence of both sgRNAs (**Fig. 1c**). Our data indicate that NHEJ damage of 5′ and 3′ overhangs was similarly robust. TA cloning confirmed that all the cleavage bands were associated with on-target mutations. However, we observed a mutant *Ar* sequence in one of 43 TA clones sequenced from cells treated with plasmids encoding Cas9(D10A) nickase and AR-A, which shows that nickases can induce double-strand breaks by nick conversion at a low efficiency[20].

Phosphorylation and recruitment of γ-H2AX (H2AX) to sites of double-strand breaks[25] can be used as a measure of global damage to the genome. We counted the number of H2AX foci in nuclei of mouse cells exposed to Cas9 or Cas9(D10A) nickase in the presence and absence of sgRNAs (**Supplementary Fig. 3** and **Supplementary Table 3**). Cells exposed to Cas9 or Cas9(D10A) alone showed very few H2AX foci (2–2.5 foci/nucleus), and Cas9 protein was localized to discrete subnuclear inclusions. In the presence of sgRNAs, Cas9 protein was distributed throughout the nucleus, and we observed prominent H2AX foci. We observed a fourfold reduction in the average number of foci in nuclei treated with vectors encoding AR-A and AR-B sgRNAs and Cas9 nickase (9 foci/nucleus) compared to Cas9 endonuclease (34 foci/nucleus). Nevertheless, the number of foci in cells treated with vectors encoding Cas9(D10A) and sgRNA was elevated compared to control cells, suggesting that there is conversion

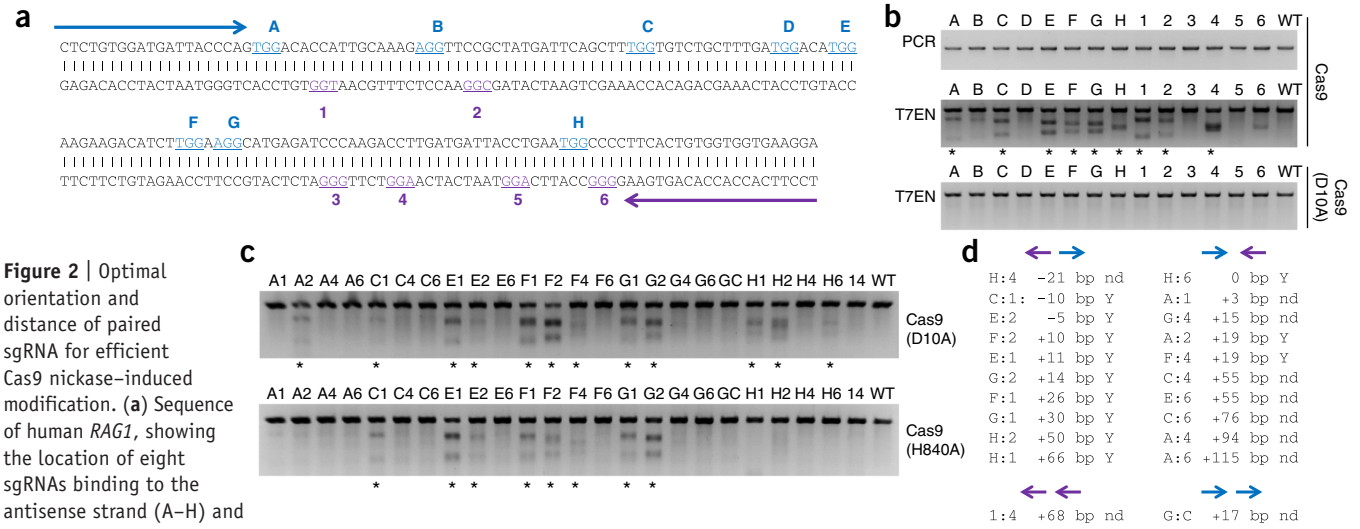

Figure 2 | Optimal orientation and distance of paired sgRNA for efficient Cas9 nickase–induced modification. (**a**) Sequence of human *RAG1*, showing the location of eight sgRNAs binding to the antisense strand (A–H) and six sites binding to the sense strand. Arrows indicate the orientation of the sgRNA sites. PAM sequence is underlined and highlighted in blue (sgRNAs A–H) and purple (sgRNAs 1–6). (**b**) T7EN cleavage assay at 14 genomic sites after co-transfection of HEK293T cells with plasmids encoding single sgRNAs and Cas9 (top) or Cas9(D10A) nickase (bottom). *, efficient cleavage. (**c**) T7EN cleavage assay demonstrating efficient cleavage (*) at 12 paired sites after co-transfection of dual sgRNAs with Cas9(D10A) nickase and 8 sites with Cas9(H840A). (**d**) Summary of T7EN cleavage data showing genome damage by Cas9 nickases (Y). The orientation of sites is indicated by arrows. Maximal activity was observed with Cas9(D10A) and Cas9(H840A) nickases at sites in a tail-to-tail orientation separated by 10–30 bp. nd, not detected. X:# indicates the sgRNA pairs shown in **a**.

of some single-strand nicks to double-strand breaks. We conclude that using Cas9 nickase will reduce, but not eliminate, potential off-target damage resulting from NHEJ at double-strand breaks.

We next co-injected one-cell embryos with either mRNA encoding Cas9(D10A) or Cas9(H840A) nickase and AR-A and AR-B sgRNAs (**Fig. 1d**). We observed no *Ar* mutations in any of the 13 founders from cells exposed to Cas9(D10A) and single sgRNA (data not shown). However, after injections of mRNA encoding Cas9(H840A) and single sgRNAs, one of 18 founders had a 48-bp deletion in *Ar*, based on the T7EN cleavage assay and sequencing (**Supplementary Fig. 4a**,**b**), which suggested that the mutant HNH domain may retain partial activity and induce double-strand breaks[17]. Co-injection of mRNA encoding Cas9 nickases and both sgRNAs yielded mutant alleles of various sizes in 13 of 20 founders (**Fig. 1d**). Sequencing of PCR products showed 25 independent alleles with insertions and/or deletions (indels) of 22–138 bp (**Fig. 1e**). To determine whether Cas9 nickases induced off-target mutations in the founder mice, we subjected PCR products for the three AR-A–AR-B off-target sites (OST1, OST3 and OST9) detected from injections of embryos with wild-type Cas9 (**Supplementary Fig. 1d**), to the T7EN cleavage assay. Within the limits of detection, we observed no off-target damage from injections of both AR-A and AR-B sgRNAs (**Fig. 1d**), nor with single sgRNAs (**Supplementary Fig. 4c**). In summary, our data show efficient cleavage of the *Ar* gene *in vivo* using Cas9 nickases and paired sgRNAs with no detectable damage at bona fide off-target sites.

To assess the specificity of the paired nickase strategy, we assayed 24 known off-target sites described for *VEGFA* (named T1–T3 in ref. 15) and *EMX1* (T4) sgRNAs in cultured human cells[15]. We selected paired sgRNAs within 50 bp of these T1–T4 sites on opposite strands in a tail-to-tail orientation (**Supplementary Fig. 5a**). In transient transfections of HeLa cells, we observed comparable on-target cleavage at all sites with Cas9 endonuclease or Cas9(D10A) nickase (**Supplementary Fig. 5b**); single sgRNAs produced no damage in the presence of the nickase, within the limits of the T7EN cleavage assay. We observed T7EN cleavage bands with wild-type Cas9 endonuclease at 20 of 24 previously identified off-target sites (**Supplementary Fig. 5c**)[15], but no damage at these 20 sites in the presence of Cas9(D10A) nickase. Deep sequencing of amplicons for each off-target region confirmed a very low frequency of indels, indistinguishable from background, in cells exposed to Cas9 nickase (**Supplementary Fig. 6** and **Supplementary Table 4**). Thus, we demonstrate that Cas9 nickase and paired sgRNAs effectively decrease off-target damage.

To define the orientation and distance between paired sites for effective damage by NHEJ, we selected nested target sites on opposing strands in exon 11 of the human *RAG1* gene (**Fig. 2a**). Based on the T7EN cleavage assay, nine of 14 sgRNAs individually induced efficient damage in the presence of Cas9 endonuclease, but not with Cas9(D10A) nickase (**Fig. 2b**) in transient transfections of human cells. Combinations of the active sgRNAs were then cotransfected with Cas9(D10A) or Cas9(H840A) nickases. PCR products for 12 paired sites exposed to Cas9(D10A) and eight paired sites exposed to Cas9(H840A) were cleaved by T7EN (**Fig. 2c**). We observed maximal cleavage at sites oriented tail-to-tail and separated by −10 bp to +30 bp (**Fig. 2d**). Finally, adjacent sites on the same DNA strand (head-to-tail orientation) did not show damage by NHEJ. While our manuscript was under review, another group[21] reported similar findings.

Next, we applied the Cas9 nickase strategy to *in vivo* modification of multiple genes. We co-injected one-cell mouse embryos with mRNA encoding Cas9(D10A) nickase and a mixture of pairs of sgRNAs to three genes; *Prkdc*, *Rag1* and *Rag2*. The paired sites for these three genes are oriented tail-to-tail and separated by 12 bp, 27 bp and 25 bp, respectively. *Rag1* and *Rag2* are 11 kbp apart on chromosome 2, which provided an opportunity to test

Table 1 | Computational analysis of paired off-target sites for mouse *Ar* sgRNAs

sgRNA pair[a]	Paired sites within 9 kb	Valid paired sites within 150 bp	Nearest paired site (bp)
AR-A–AR-B	10	1	109
AR-A–AR-A	31	0	828
AR-B–AR-B	6	1	−10

[a]The number of potential off-target sites for individual sgRNAs, AR-A and AR-B, was 2,821 and 1,027, respectively.

the efficacy of Cas9 nickase for the generation of sizeable deletions of genomic DNA, as described for wild-type Cas9 (ref. 14). After injection of 175 embryos, we recovered 15 founder mice and assayed them for NHEJ-induced mutations in the three target genes (**Supplementary Fig. 7**). We recovered six mutant alleles for *Prkdc*, five mutant alleles for *Rag1* and five mutant alleles for *Rag2*. One founder mouse contained mutant alleles for all three genes (founder 78). Furthermore, we detected deletions of the region between *Rag1* and *Rag2* in 8 founder mice. Our results demonstrate that a multiplex paired nickase strategy is effective for modification of multiple genes *in vivo*, as previously reported with wild-type Cas9 endonuclease[13]. Our findings also show that the paired nickase strategy can be deployed to generate sizeable deletions in the genome at a high efficiency. Finally, we observed no detectable damage in the founder animals at two off-target sites for *Prkdc* and *Rag1* that are susceptible to *in vivo* cleavage by wild-type Cas9 (**Supplementary Fig. 8** and **Supplementary Table 5**).

Our experiments established a framework for the optimal design of paired sgRNAs and rules for the computation of potential off-target sites for sgRNA pairs. Cas9 sites in a tail-to-tail orientation and optimally separated by −10 bp to +30 bp will have the sequence 5′-CCN$_{(32–72)}$GG. This motif is very common in protein-coding exons (http://www.sanger.ac.uk/htgt/wge/). For example, in the mouse exome, we found 7,764,364 paired sites covering 95% of all coding exons (**Supplementary Fig. 9**), with an average of 342 paired sites per gene. In contrast, potential off-target sites for any pair of Cas9 sites are expected to be rare as similar sequences are unlikely to occur close together elsewhere in the genome. Allowing for up to five mismatches in the sgRNA sequence, we found only two possible paired sites within 150 bp for the AR-A–AR-B sgRNAs (**Table 1** and **Supplementary Data**). We detected no NHEJ damage at these two potential paired off-target sites in the founder animals (**Supplementary Fig. 10**). Furthermore, computational analysis of all possible combinations of the *Prkdc*, *Rag1* and *Rag2* sgRNAs did not reveal any paired off-target sites that mapped within 150 bp (**Supplementary Table 6**).

In summary, the use of Cas9 nickase and paired sgRNAs greatly increased the fidelity of the Cas9 system for *in vivo* genome editing without compromising efficiency. In both mouse embryos and in cultured cells, we found no evidence for NHEJ-induced damage by Cas9 nickase at off-target sites prone to mutation by wild-type Cas9 endonuclease. The strategy can be applied on a genome-wide scale to any model organism or model cell and is particularly relevant for therapeutic applications where off-target damage is a concern.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank members of the entire Huang laboratory for their support and advice, G. Tischler for the C++ alignment code, and J. Liu and C. Gao for help with deep sequencing. This work was supported by grants from the 973 program (2010CB945101 and 2011CB944301) and a core grant from the Wellcome Trust.

AUTHOR CONTRIBUTIONS

B.S., W.Z., J. Zhang, X.H. and W.C.S. designed the experiments and wrote the manuscript. B.S., W.Z., J. Zhang, J. Zhou, J.W. and L.C. performed the experiments. L.W., A.H. and V.I. performed the computational analysis of off-target sites.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. *Nat. Rev. Genet.* **11**, 636–646 (2010).
2. Joung, J.K. & Sander, J.D. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
3. Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. *Nature* **482**, 331–338 (2012).
4. Jinek, M. *et al. Science* **337**, 816–821 (2012).
5. Cong, L. *et al. Science* **339**, 819–823 (2013).
6. Mali, P. *et al. Science* **339**, 823–826 (2013).
7. Hwang, W.Y. *et al. Nat. Biotechnol.* **31**, 227–229 (2013).
8. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. *Nat. Biotechnol.* **31**, 230–232 (2013).
9. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. *Nat. Biotechnol.* **31**, 233–239 (2013).
10. Jinek, M. *et al. Elife* **2**, e00471 (2013).
11. Shen, B. *et al. Cell Res.* **23**, 720–723 (2013).
12. Qi, L.S. *et al. Cell* **152**, 1173–1183 (2013).
13. Wang, H. *et al. Cell* **153**, 910–918 (2013).
14. Yang, H. *et al. Cell* **154**, 1370–1379 (2013).
15. Fu, Y. *et al. Nat. Biotechnol.* **31**, 822–826 (2013).
16. Hsu, P.D. *et al. Nat. Biotechnol.* **31**, 827–832 (2013).
17. Mali, P. *et al. Nat. Biotechnol.* **31**, 833–838 (2013).
18. Pattanayak, V. *et al. Nat. Biotechnol.* **31**, 839–843 (2013).
19. Wang, J. *et al. Genome Res.* **22**, 1316–1326 (2012).
20. Kim, E. *et al. Genome Res.* **22**, 1327–1333 (2012).
21. Ran, F.A. *et al. Cell* **154**, 1380–1389 (2013).
22. Yeh, S. *et al. Proc. Natl. Acad. Sci. USA* **99**, 13498–13503 (2002).
23. Tesson, L. *et al. Nat. Biotechnol.* **29**, 695–696 (2011).
24. Caldecott, K.W. *Nat. Rev. Genet.* **9**, 619–631 (2008).
25. Rogakou, E.P., Boon, C., Redon, C. & Bonner, W.M. *J. Cell Biol.* **146**, 905–916 (1999).

ONLINE METHODS

Animals. Mice were housed in standard cages in an Assessment and Accreditation of Laboratory Animal Care–credited specific pathogen–free (SPF) animal facility on a 12-h light–12 h dark cycle. All animal protocols were approved by the Animal Care and Use Committee of the Model Animal Research Center, the host for the National Resource Center for Mutant Mice in China, Nanjing University. For the animal studies, randomization was limited as the animals were not assigned to groups for treatments. The phenotyping studies were randomized with time. All other sources of variation were managed by standardization. Investigators were not blinded to the founder mouse generation profile as cage cards displayed this information. Instead, unbiased characterization relied on randomization and standardization.

DNA constructs. The Cas9 expression construct pST1374-Cas9-N-NLS-Flag-linker (Addgene 44758) was optimized for nuclear import and contains a CMV promoter for expression in mammalian cells and a T7 promoter for *in vitro* transcription of *Cas9* mRNA[11]. This plasmid also contains a blasticidin selection cassette for transient enrichment of transfected cells expressing Cas9. To generate the vectors encoding Cas9(D10A) and Cas9(H840A) nickases, pST1374-Cas9-N-NLS-flag-linker was mutagenized by PCR amplification using PrimeSTAR HS DNA polymerase (Takara, DR010A) and the In-Fusion PCR Cloning kit (Clontech, 638909). The primers used for site-specific mutagenesis are Mut D10A forward (For): 5′-GGACTGGCTATCGGGACAAACTCC GTTGGCTGGGC; Mut D10A reverse (Rev): 5′-CCCGATAGC CAGTCCAATAGAGTATTTCTTGTCTCTAG; Mut H840A For: 5′-CGTGGACGCTATTGTTCCACAGTCCTTCCTCAAAGATG; and Mut H840A Rev: 5′- AACAATAGCGTCCACGTCATAATCG GACAGCCGGTTG. The pUC57-sgRNA expression vector used for *in vitro* transcription of sgRNAs contains a synthetic fragment composed of a T7 promoter, two BsaI sites and the sgRNA cloned into the pUC57 vector (Thermo, SD0171; modified by Taihe, Beijing). pGL3-U6-sgRNA-PGK-Puro vector used for expression of sgRNAs in cells contains the U6-PGK-Puro fragment amplified from pLKO.1 (Addgene 8453) inserted into pGL3-Basic plasmid backbone (Promega, E1751). The sgRNA scaffold was PCR-amplified from pUC57-sgRNA and inserted downstream of *U6* promoter using AgeI and EcoRI. Oligos for the generation of sgRNA expression plasmids (**Supplementary Table 7**) were annealed and cloned into the BsaI sites of pUC57-sgRNA or pGL3-U6-sgRNA-PGK-Puro. Annotated sequences of pST1374-Cas9-N-NLS-flag-linker-D10A (Addgene 51130), pST1374-Cas9-N-NLS-flag-linker-H840A (Addgene 51129), pUC57-sgRNA (Addgene 51132) and pGL3-U6-sgRNA-PGK-Puro (Addgene 51133) plasmids are presented in the **Supplementary Note**.

***In vitro* transcription.** Cas9 expression plasmid was linearized with Age1 and *in vitro* transcribed using the T7 Ultra Kit (Ambion, AM1345). mRNA was purified by RNeasy Mini Kit (Qiagen, 74104). pUC57-sgRNA expression vectors were linearized by DraI and *in vitro*–transcribed using the MEGAshortscript Kit (Ambion, AM1354). sgRNAs were purified by MEGAclear Kit (Ambion, AM1908) and recovered by alcohol precipitation.

***Cas9* mRNA–sgRNA injection of one-cell embryos.** *Cas9* and sgRNA injections of one-cell embryos were performed as described[11]. Briefly, mouse zygotes obtained by mating of CBA males with superovulated C57BL/6J females were injected with a mixture of *NLS-flag-linker-Cas9* mRNA and sgRNAs as indicated in **Supplementary Table 8**. Microinjections were performed into the cytoplasm and larger (male) pronucleus of fertilized oocytes. Injected zygotes were transferred into pseudo-pregnant CD1 female mice from which viable founder animals were obtained. Sufficient injections were completed to ensure that adequate founder animals would be available for characterization. All animals produced were used, the exact number varied depending on the efficiency of the process.

Cell culture and electroporation. NIH3T3, HEK293T and HeLa cells were cultured in DMEM–high glucose (HyClone, SH30022.01B) with 10% FBS, penicillin (100 U/ml) and streptomycin (100 μg/ml). 2×10^6 cells were electroporated (Bio-Rad Gene Pulser XL) with 4 μg of *Cas9* expression plasmids and 4 μg of pGL3-U6-sgRNA-PGK-Puro. For co-transfection of paired sgRNAs, 2 μg of each sgRNA plasmid were used. Empty pGL3-U6-sgRNA-PGK-Puro plasmid was used as control. Puromycin (1 μg/ml, Merck, 540411) and blasticidin (10 μg/ml, Sigma, 15205) were added 24 h after transfection. Cells were collected 72 h after transfection.

T7EN cleavage assay. Cells and mouse tail biopsies were collected and digested with 100 μg/ml Proteinase K in lysis buffer (10 μM Tris-HCl, 0.4 M NaCl, 2 μM EDTA and 1% SDS). Genomic DNA was extracted by phenol-chloroform and alcohol precipitation. The T7EN cleavage assay was performed as described[11]. In brief, targeted regions were PCR-amplified from genomic DNA and the products were purified with a PCR cleanup kit (Axygen, AP-PCR-50). Purified PCR product was denatured and reannealed in NEBuffer 2 (NEB) using a thermocycler. Hybridized PCR products were digested with T7EN (NEB, M0302L) for 30 min and separated by 2% agarose gel. To detect T7EN cleavage products in *Ar*-deficient mice, 50 ng of PCR fragment from wild-type mice was mixed with 150 ng of PCR fragments from founder mice (F1–F20, F41–F71). Primer pairs for PCR amplification of mouse *Ar*, *Prkdc*, *Rag1*, *Rag2* and human *RAG1* target regions are listed in **Supplementary Table 9**. Primer pairs for PCR amplification of off-target regions for *Ar*, *Prkdc* and *Rag1* sgRNAs are listed in **Supplementary Tables 10–12**. Primer pairs for PCR amplification of *Ar* paired off-target sites are listed in **Supplementary Table 13**.

H2AX staining and analysis. 2×10^5 3T3 cells were plated on slides in 6-well plates and transfected 24 h later with Lipofectamine LTX (Invitrogen, 15338-100) containing 2 μg of Cas9 plasmid, 2 μg Cas9(D10A) plasmid, 2 μg Cas9 and 2 μg AR-A + 2 μg AR-B sgRNAs, or 2 μg Cas9(D10A) plasmid and 2 μg AR-A + 2 μg AR-B sgRNAs. 48 h after transfection, cells were fixed and stained with anti–γ-H2AX (CST, 9718) and anti-Flag (Sigma, F1804) antibodies, followed by incubation with FITC-conjugated goat anti-mouse IgG (Jackson, 115-095-146), Cy3-conjugated goat anti-rabbit IgG (Jackson, 111-165-003) and DAPI. All fluorescence images were taken on an Olympus Fluoview 1000 confocal microscope using the same settings. H2AX and Flag double-positive nuclei were counted (30 cells for each group randomly selected, **Supplementary Table 3**). A two-sided student's

t-test was used to evaluate the difference in number of foci between experimental groups; with a sample size of 30 this test has a statistical power to detect a large effect (as defined by Cohen's *d* as 0.8) the majority of the time (power = 0.85). This statistical method is robust, as the data were found to be normally distributed and to have equal variance between the groups being compared. Differences were considered to be significant at $P < 0.05$.

Deep sequencing of off-target sites. Off-target loci were amplified by primers listed in **Supplementary Table 14**. PCR products were purified by Axygen PCR Cleanup kit and combined into separate equimolar pools for the control, Cas9 and Cas9(D10A) samples. Library preparation was performed according to the manufacturer's instructions (Ion Torrent). Ion Xpress barcoded adaptors 053, 054 and 055 were used for the control, Cas9 and Cas9(D10A) libraries, respectively. After emulsion PCR, the three libraries were multiplexed on an Ion 316 chip at a 1:1:1 mass ratio to maintain equimolar amounts of each amplicon. Sequencing was performed in a PGM sequencer (Ion Torrent) using the Ion PGM sequencing 400 kit according to the manufacturer's instructions. Variants were identified with custom software. Indels of three or more base pairs located around the sgRNA target site were considered to be NHEJ-mediated modifications.

sgRNA design and identification of off-target sites. For the mouse *Ar* gene, the AR-A and AR-B Cas9 sites were selected with the sequence 5′-GGN$_{(19)}$GG (**Supplementary Table 7**). The 5′ dinucleotide GG ensures optimal expression from the *T7* and *U6* promoters. Highly similar sequences in the genome were identified by BLAST and a random set of 8 potential off-target sites were selected for AR-A and AR-B (**Supplementary Fig. 1c**). For human genes, Cas9 sites were selected with the sequence 5′-N$_{(21)}$GG. The 5′ end of the transcribed sgRNA sequence was appended with a G or GG nucleotide for optimal expression (**Supplementary Table 7**). Putative off-target sites in the mouse or human genome were identified as follows. All possible sites with homology to the 23 bp sequence (sgRNA + PAM) were retrieved by a base-by-base scan of the entire mouse genome (C++ code available upon request), allowing for ungapped alignments with up to 5 mismatches in the sgRNA target sequence. The scan is made suitably efficient by streaming the genome sequence, and using bitwise (rapid) operations to check for the number of mismatches between the (sgRNA + PAM) sequence and the specific segments of the genome. The output of this scan is in bed format[26]. Off-target sites for a set of paired Cas9 sites are identified from the genomic coordinates of all off-target target sites computed for individual Cas9 sites. BedTools[26] utilities are used to find all possible combinations of sites in a tail-to-tail or head-to-head orientation within 9 kb on the same chromosome and to calculate the distance between the sites. The ten closest sites for each combination and orientation of mouse *Ar* Cas9 sites (AR-A and AR-B) is shown in **Supplementary Data**.

26. Quinlan, A.R. & Hall, I.M. *Bioinformatics* **26**, 841–842 (2010).

doi:10.1038/nmeth.2857